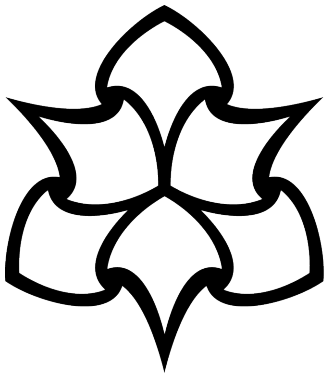


Three clinically distinct chronic pediatric airway infections share a common core microbiota

Running head: Common pediatric core respiratory microbiota

Christopher J. van der Gast[1], Leah Cuthbertson[1], Geraint B. Rogers[2], Christopher Pope[3,4], Robyn L. Marsh[5], Gregory J. Redding[3], Kenneth D. Bruce[6], Anne B. Chang[5,7], Lucas R. Hoffman[3,4]*

[1] *NERC Centre for Ecology & Hydrology, Wallingford, OX10 8BB, UK*

[2] *Immunity, Infection and Inflammation Program, Mater Research, Translational Research Institute, South Brisbane, QLD, 4101, Australia*

Departments of [3]*Pediatrics and* [4]*Microbiology, University of Washington, Box 356320, HSB K328A, Seattle, WA 98105, USA*

[5] *Menzies School of Health Research, Charles Darwin University, Darwin, NT, 0810, Australia*

[6] *Institute of Pharmaceutical Science, Molecular Microbiology Research Laboratory, King's College London, London, SE1 9NH, UK*

[7] *Queensland Children's Respiratory Centre, Royal Children's Hospital, Herston, Brisbane, QLD, 4029, Australia*

* For correspondence: Address: 4800 Sand Point Way NE, Box A-5937, Seattle Children's Hospital, Seattle, WA, 98105. E-mail: lhoffm@u.washington.edu.

Word count: Abstract, 328 words. Text, 4,028 words.

Portions of this work were presented at the 2013 European Cystic Fibrosis Society Conference in Lisbon, Portugal, June, 2013.

Descriptor: 10.11 Pediatrics: Respiratory Infections

Author contributions: CJvdG designed and performed the study and wrote the manuscript. LC, GBR, CP, and RLM performed analyses and wrote the manuscript. KDB and GJR designed the study and wrote the

manuscript. ABC and LRH designed the study, contributed specimens, performed the study, and wrote the manuscript.

This study was supported by grants from the Cystic Fibrosis Foundation (HOFFMA07P0) and the American Thoracic Society (CF-07-003), the NIH (K02HL105543), the UK Natural Environment Research Council (NE/H019456/1), and Australia's National Health and Medical Research Council grants (1042601, 1040830, 1019834 and 1034703).

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

Abstract:

Rationale: DNA-based microbiological studies are moving beyond studying healthy human microbiota to investigate diverse infectious diseases, including chronic respiratory infections such as those in the airways of people with cystic fibrosis (CF) and non-CF bronchiectasis (BE). The species identified in the respiratory secretion microbiota from such patients can be classified into those that are common and abundant among similar subjects (core) versus those that are infrequent and rare (satellite). This categorization provides a vital foundation for investigating disease pathogenesis and improving therapy. However, whether the core microbiota of people with different respiratory diseases, which are traditionally associated with specific culturable pathogens, are unique or shared with other chronic infections of the lower airways is not well studied. Little is also known about how these chronic infection microbiota change from childhood into adulthood.

Objectives: We sought to compare the core microbiota in respiratory specimens from children and adults with different chronic lung infections.

Methods: We used bacterial 16S rRNA gene pyrosequencing, phylogenetic analysis, and ecological statistical tools to compare the core microbiota in respiratory samples from three cohorts of symptomatic children with clinically distinct airway diseases (protracted bacterial bronchitis, BE, CF), and four healthy children. We then compared the core pediatric respiratory microbiota with those in samples from adults with BE and CF.

Measurements and Main Results: All three pediatric disease cohorts shared strikingly similar core respiratory microbiota that differed from adult CF and BE microbiota. The most common species in pediatric disease cohort samples were also detected in those from healthy children. The adult CF and BE microbiota also differed from each other, suggesting common early infection airway microbiota that diverge by adulthood. The shared core pediatric microbiota included both traditional pathogens and many species not routinely identified by standard culture.

Conclusions: Our results indicate that these clinically distinct chronic airway infections share common early core microbiota, which are likely shaped by natural aspiration and impaired clearance of the same airway microbes, but that disease-specific characteristics select for divergent microbiota by adulthood. Longitudinal and

interventional studies will be required to define the relationships between microbiota, treatments, and disease progression.

Number of words in this abstract: 349

Introduction

A growing body of evidence indicates that the airways of even healthy people contain detectable microbes (1–4), and that these microbiota are altered in many chronic lung diseases. For example, culture-independent analyses of cystic fibrosis (CF) respiratory samples have revealed remarkably diverse microbial communities compared with culture results. Among the bacteria detected most frequently, and at highest abundance, in those CF samples were anaerobic species not usually detected by standard laboratory methods (5). Both cross-sectional and longitudinal studies indicate that CF airway microbiota generally decrease in diversity over time, concurrent with a decrease in average lung function (6–9). Whether these ecological changes are the cause of lung function decline, or whether they are the result of increasing antibiotic exposure or changing properties of the infected airways, are not understood (9). Furthermore, exactly where the species that comprise CF secretion microbiota come from, and whether features unique to the CF airway or its secretions select for CF-specific microbiota, are unknown.

Several other childhood lung diseases also involve chronic airway infection, including non-CF bronchiectasis (BE) and protracted bacterial bronchitis (PBB) (10). At least some underlying factors that predispose to infection in CF, BE, and PBB may be shared (e.g., mucus hypersecretion or stasis, impaired airway clearance), while others vary (e.g., inflammatory response, intensity of antibiotic use, physicochemical mucus characteristics), and the associated infectious microbiota could therefore differ at the beginning or throughout these diseases. In support of this concept, specific microbes have been associated with the pathogenesis of each of these diseases in culture-based studies. For example, *Staphylococcus aureus* and *Pseudomonas aeruginosa* are traditionally associated with CF lung disease (11). In contrast, *Streptococcus pneumoniae*, *Moraxella catarrhalis*, and *Haemophilus influenzae* are traditionally associated with both PBB and early BE (10), with *P. aeruginosa* associated with more advanced BE disease (12). Treatment regimens therefore tend to differ for each disease, designed to target specific pathogens. Moreover, it has been suggested that these bacteria play primary roles in initiating or driving associated symptoms (13), which could imply disease-specific infection pathogenesis. However, it is now known that children with CF have lung disease without detectable infection by *P. aeruginosa* or *S. aureus* (14, 15). The identification of diverse microbiota in even healthy airways further complicates the task of delineating “infected” and/or “pathogenic” from “normal” conditions. Therefore, comparing the earliest stages of chronic infection for diverse airways diseases could be helpful in understanding whether infection begins (and drives disease) in distinct or similar ways in different conditions.

An analytical approach that has been particularly useful in understanding environmental and infectious microbiota involves partitioning the constituent microbes into core and satellite species (7). In cross-sectional studies, core species are those that are detected in the majority of samples from subjects with a specific disease, while satellite species are those that are infrequent. While this categorization does not necessarily identify which species are pathogens, it provides a useful framework within which to compare microbiota from different subject groups, and the resulting similarities and differences have important implications for infection pathogenesis. Here, we compared the core and satellite microbiota in three very different cohorts of children with PBB, BE, and CF using next-generation sequencing and ecological analytical approaches. Because these study subjects were children, and therefore at relatively early stages of disease, we hypothesized that these cohorts would have similar, rather than disease-specific, early core respiratory sample microbiota. We then compared pediatric and adult CF and BE respiratory metacommunities, using data from two of our recent adult studies (16, 17) to identify evidence of disease-specific changes in microbiota as patients with each disease progress to adulthood.

Materials and methods

Patients and clinical samples

All children in this study were participants in previous studies of CF, PBB, and BE, respectively (18–20). Inclusion and exclusion criteria and study details are listed in the Online Supplement. Sputum and bronchoalveolar lavage (BAL) fluid, as well as clinical data, were collected during those studies as described (18–20). The study was approved by the institutional review boards at Seattle Children's Hospital (SCH) for the CF samples, and at the Royal Children's Hospital (RCH, Brisbane, Australia) for the PBB, BE and control pediatric samples. Data from two adult BE (n = 38 samples, age 37 - 74 years) and CF (n = 30, 18 - 55 years) cohorts from previous studies were used for metacommunity comparisons with the microbiota from the pediatric cohorts (16, 17).

DNA extraction and Q-PCR

Full details of DNA extraction from BAL and sputum samples, and of the use of quantitative PCR to determine total bacterial abundance using universal eubacterial primers, are in the Online Supplement.

Bacterial Tag-Encoded FLX Amplicon Pyrosequencing (bTEFAP) and sequence analyses

The bacterial taxonomy of each sample was evaluated using bacterial tag-encoded FLX-titanium pyrosequencing (bTEFAP) of a fragment of the 16S rRNA gene covering the V1-V3 variable regions, as described previously (21). Details of amplification, sequencing, and sequence processing and analysis are in the Online Supplement.

Statistical analyses

To avoid potential biases in comparisons of diversity between local communities due to varying number of sequences per sample, a randomised re-sampling method, using three indices of diversity (S*, H', and 1-D), were employed as previously described (7). Bacterial species within each metacommunity were partitioned into core and satellite species groups using the Poisson distribution test as previously described (7) (Online Supplement).

Results

Pediatric study cohorts

We analysed the microbiota in respiratory samples (sputum and BAL fluid) from 3 cohorts of children with BE, PBB, and CF, respectively. All of these children either underwent bronchoscopy for chronic wet cough or were expectorating, indicative of symptomatic disease. As shown in Table 1, the cohorts were substantially different in terms of demographic and treatment features that could play important roles in determining the airway microbiota, including age, antibiotic exposure, and *P. aeruginosa* culture-positivity, which have each been shown to correlate inversely with airway microbiota diversity in CF, and FEV_1% predicted, which correlates directly with diversity (9). The children with CF also had higher rates of inhaled corticosteroid use, a treatment that could conceivably alter airway microbiota constituency (2). The CF cohort originated entirely from a single hospital in the US, whereas all children with BE and PBB were from a single hospital in Australia. Each of these differences, if they were to influence airway microbiota, would be unlikely to bias our analyses towards finding microbial similarities between the three cohorts.

Pediatric community diversity and composition

Bacterial pyrosequencing data were used to assess and compare diversity and composition of microbiota from respiratory samples within and across the 3 pediatric cohorts. A total of 408,335 bacterial sequence reads (mean ± SD per sample, 5041 ± 3557), identifying 104 genera and 225 distinct operational taxonomic units (OTUs) classified to species level (Table S1), were generated from all samples combined. The average numbers of bacterial sequence reads per sample were similar among the three cohorts (mean ± SD: BE, 4284 ± 3476; PBB, 4471 ± 3718; and CF, 4785 ± 2702). Estimates of total bacterial abundances were also similar when quantified by qPCR (for all samples, including sputum and BAL cell pellets, from which sufficient DNA remained after sequencing analysis, mean cfu/mL equivalents/sample ± SD: BE, $1.47 \times 10^7 \pm 2.32 \times 10^7$ *N*=18 samples; PBB, $9.89 \times 10^6 \pm 1.38 \times 10^7$, *N*=10; and CF, $5.82 \times 10^7 \pm 6.28 \times 10^7$, *N*=21).

Bacterial diversity within and between cohorts was compared using three indices of diversity: species richness (*S**, the total number of species), Shannon-Wiener (*H*', a metric that accounts for both number and relative abundance of species), and Simpson's (1-*D*, a measure of the probability that two species randomly selected from a sample will differ). Bacterial diversity between BE sputum and BAL samples was compared using all three measures and found not to be significantly different ($P > 0.05$, Fig. S1), and hence those samples

were combined for subsequent analyses. S^* was found to vary substantially within cohorts, with means of 30.6 ± 10.5, 30.9 ± 11.5, and 19.5 ± 10.0 for the BE, PBB, and CF cohorts, respectively (Fig. 1); similar results were obtained for H' and 1-D (Fig. 1 and Table S2). Diversity between pediatric cohorts was not significantly different ($P > 0.05$) in all instances except for one measure (H') between the PBB and CF cohorts (Table S2).

Similarities and differences in community membership and structure for microbiota samples both within and between disease cohorts were assayed with two different indices: Sørensen (S_{SOR}, which accounts for the number of species present in each community and those that are shared) and Bray-Curtis quantitative (S_{BC}, similar to S_{SOR} but also accounts for abundance of each species). Each index varies in value from 0 to 1, with higher values indicating greater similarity. It has been previously observed that community composition is highly variable between CF patients (7, 22). Here, we found that observation held regardless of underlying pediatric lung disease, with mean similarities in bacterial community membership taken pair-wise for samples within pediatric cohorts of $S_{SOR} = 0.44 \pm 0.14$ for BE ($n = 171$ pair wise comparisons); $S_{SOR} = 0.51 \pm 0.12$ for PBB ($n = 66$); and $S_{SOR} = 0.50 \pm 0.12$ for CF ($n = 300$), with a mean similarity of the entire pool of samples of $S_{SOR} = 0.46 \pm 0.13$ ($n = 1540$) (Fig. 2). For S_{BC}, the mean between all samples was $S_{BC} = 0.19 \pm 0.15$ ($n = 1540$), while that within cohorts was $S_{BC} = 0.18 \pm 0.14$ for BE ($n = 171$); $S_{BC} = 0.27 \pm 0.12$ for PBB ($n = 66$); and $S_{BC} = 0.27 \pm 0.17$ for CF ($n = 300$). Therefore, individual sample microbiota differed within each pediatric disease group to a similar degree, and the microbiota identified within each pediatric cohort's samples were pooled together into disease-specific metacommunities (defined as the set of all microbiota identified in samples from a given cohort) for subsequent comparisons.

Microbiota within a metacommunity framework

Previously, we established that the categorization of component species in CF microbiota into core and satellite species revealed important aspects of species-abundance distributions within a metacommunity that would be neglected without such a distinction (7). A coherent metacommunity could be expected to exhibit a direct relationship between prevalence and abundance of individual species within the constituent communities. Consistent with this prediction, the abundance of species in each pediatric study cohort was significantly correlated with the number of individual respiratory sample communities those species occupied (Fig. 3A).

The species identified in each pediatric cohort metacommunity were then classified as core or satellite based upon their distributions (Fig. 3B). Of the 177 species that comprised the BE metacommunity, 88 were core

and 89 were satellite species. The PBB metacommunity (130 species) was comprised of 63 core and 67 satellite species, and the CF metacommunity (143 species) comprised of 66 and 77 species, respectively.

Similarities between the three pediatric metacommunities were assayed using S_{SOR} and S_{BC} indices for all species, and for the core and satellite groups. The resulting cluster analysis revealed community membership to be relatively well-conserved between cohorts for all species (S_{SOR} = 0.74 ± 0.02 [n = 3 pair-wise comparisons]), but much more highly conserved between the core groups (S_{SOR} = 0.93 ± 0.02) (Fig. 4). The satellite groups, comprised of randomly distributed species, were highly divergent (S_{SOR} = 0.35 ± 0.01). Results from Bray-Curtis quantitative analyses were similar: All S_{BC} = 0.50 ± 0.10; core S_{BC} = 0.51 ± 0.10; and satellite S_{BC} = 0.10 ± 0.02. These results indicate that the high level of similarity between the three pediatric cohort metacommunities is attributable to core species.

To determine whether disease-specific microbiota emerge with disease progression into adulthood for the two chronic diseases, BE and CF (PBB usually improves long-term with antibiotic treatment (10)), we compared the microbiota from our previous studies of adult CF (16) and BE (17) cohorts with the pediatric cohort microbiota. The adult BE cohort metacommunity (n = 38 patients) was comprised of 86 core and 54 satellite species, while the adult CF cohort (n = 30) metacommunity was comprised of 68 core and 81 satellite species. The resulting analyses revealed adult community membership to be highly divergent in all cases (whole microbiota S_{SOR} = 0.36 ± 0.03; core S_{SOR} = 0.46 ± 0.06; and satellite S_{SOR} = 0.09 ± 0.03 [n = 6 pairwise comparisons in all cases]) (Fig. 4). When relative abundances of constituent taxa were included to examine differences in community structure, an even more pronounced divergence was observed than for membership (whole S_{BC} = 0.07 ± 0.03; core S_{BC} = 0.07 ± 0.04; and satellite S_{BC} = 0.01 ± 0.01 [n = 6 in all cases]). In addition, comparisons between the adult CF and BE metacommunities also revealed a high degree of divergence both in terms of membership (whole S_{SOR} = 0.38; core S_{SOR} = 0.48; and satellite S_{SOR} = 0.16) (Fig. 4) and structure (whole S_{BC} = 0.05; core S_{BC} = 0.05; and satellite S_{BC} = 0.02). Therefore, the adult CF and BE microbiota differed substantially from each other, contrasting with the similarities among the pediatric cohort microbiota.

To test whether the adult and pediatric metacommunities were significantly similar or dissimilar, the whole, core, and satellite microbiota were compared between cohorts using the Raup and Crick probability-based similarity index (S_{RC}) to determine whether compositional similarities in the microbiota were more or less significantly similar than expected by chance (Fig. 5). The resulting cluster analysis revealed the pediatric whole, core, and satellite microbiota to be significantly dissimilar (S_{RC} < 0.05) from the corresponding adult CF and BE

microbiota. Likewise, the whole and core microbiota from the pediatric cohorts were significantly similar (S_{RC} > 0.95), adding further weight to the observations of pediatric community membership (S_{SOR}) and structure (S_{BC}) (Fig. 5). In contrast, the pediatric satellite microbiota were not significantly similar or dissimilar to each other (S_{RC} > 0.05 and < 0.95), which would be expected given that satellite microbiota in a given cohort represent rare and randomly distributed species. Furthermore, the adult CF and BE microbiota were significantly dissimilar to each other (S_{RC} < 0.05) (Fig. 5).

Similarity of percentages (SIMPER) analysis of the metacommunities was used to identify those species that contributed most to the observed similarity between the three pediatric cohorts. These species, all core group members, are listed in decreasing order of contribution in Table 2. *H. influenzae* contributed the greatest amount to the observed similarity between pediatric samples. Most of the major contributors have previously been associated with respiratory tract infections and other opportunistic infections, and many are known inhabitants of the oral cavity, including both aerobic and anaerobic species (Table S1). These findings contrasted with the SIMPER analysis of the adult CF cohort, wherein *P. aeruginosa* contributed over 83% to the observed similarity between samples (Table S3a) in addition to 3 other species: *Streptococcus pneumoniae*, *Prevotella melaninogenica*, and *Veillonella parvula*. Within the adult BE cohort, the main contributors to the overall similarity between samples were *H. influenza*e (50.4%), *P. aeruginosa* (8.62 %) and *V. dispar* (8.11%) (Table S3b) along with *S. pneumoniae* and, as for the pediatric cohorts, other species associated with respiratory infections, many of which are known oral microbiota members. As also observed for the pediatric cohorts, the species identified from the SIMPER analyses for both adult cohorts were all core species in their respective metacommunities. Although there were some species identified in common for the adult and pediatric cohorts by the SIMPER analyses, these similarities were greatly outweighed by the differences in species content, accounting for the observed differences in whole and core microbiota within each adult-pediatric disease cohort pair.

To investigate whether the core microbiota shared between the three pediatric disease groups were unique to children with airways disease, we identified the microbiota in BALF from four children who underwent bronchoscopy for reasons other than suspected infection (described in the Online Supplement and [16]). Of the 10 shared core species that contributed the most to similarity between disease cohorts from Table 2, six were detected in all four “control” samples, and the other four species were detected in three of the four samples, suggesting that the shared core microbiota are commonly found in healthy, as well as diseased, children’s

airways (Table S4), consistent with prior airway microbiota studies that included samples from healthy subjects[3,21].

Discussion

We found that cohorts of symptomatic children with three clinically distinct airways diseases- CF, BE, and PBB- shared remarkably similar respiratory sample core microbiota, the most common of which were also found in airway samples from children without clinical evidence of infection. The shared core microbiota included both traditional pathogens and many bacteria that are either not identified by routine clinical laboratory methods or are generally classified as “oral flora” and are routinely ignored with respect to pathophysiology or treatment. In contrast, the respiratory sample microbiota from adults with CF and BE differed significantly from each other and from those of children with the same disease. These results indicate that the core microbes in different chronic respiratory infection types likely begin similarly, yet are divergent by adulthood.

While these observations provide valuable insight into the microbiota in diseased airways, they also identify next steps to take in understanding the pathogenesis of diverse airway infections. Given the high abundance and prevalence of the core species among pediatric patients, one interpretation of these results could be that the core species represent those that contribute the most to airway inflammation and disease progression. Alternatively, it could be that the satellite species play key pathogenic roles in specific contexts or stages of infection; for example, infection at any stage with specific low-abundance or infrequent bacteria could lead to increased inflammation, airway damage, mucus hypersecretion, and/or clinical deterioration. While these questions are not fully answered by the current study, the relative rarity of any specific species in each satellite group may indicate a consistent, general role for the core groups in pathogenesis. Future studies will be required to longitudinally compare the changes over time in airway microbiota in health and disease to better understand the microbial contribution to pathogenesis, and the best therapeutic approach to infection.

These findings provide valuable perspective for results from studies of many respiratory diseases, usually at more advanced stages (2). A study using a relatively low-depth sequencing technique, 16S rRNA gene clone sequencing, comparing microbiota in explanted lungs from adults with end-stage BE and CF identified diverse communities often dominated by *Pseudomonas* in both sample sets (23). Likewise, two studies using the same technique applied to sputa from patients with CF and BE found some similarities and some differences between the subject groups, but no formal comparisons were made (24, 25). In support of our results (17), a recent pyrosequencing study of 21 older adults with non-CF bronchiectasis showed their sputum to contain diverse microbiota (26). Studies of sputum, BAL fluid, and explanted lung tissue from patients at various stages of disease due to chronic obstructive pulmonary disease (COPD) and asthma, sometimes comparing with

297 healthy subjects, have yielded varying and often conflicting results, in some cases demonstrating disease-
298 specific microbiota and in others identifying substantial overlap (1, 2, 27–29). In all of these studies, the effects
299 of therapies that could conceivably impact airway microbiota, such as intubation, steroids and antibiotics, were
300 difficult to control for. Our results suggest that the microbiota observed in those older patients, at later stages of
301 disease, likely had similar (or identical) beginnings and diverged over time. Whether those changes resulted
302 from features of the underlying disease or from differences in treatment are not yet clear. While the differences in
303 antibiotic use among study cohorts support the idea that antibiotics did not have a profound, lasting impact on
304 airway microbiota (similar to observations from CF (8, 30)), and there was no evidence that differences in steroid
305 use resulted in differences in microbiota in these children, chronic or repetitive drug exposure could have
306 gradual effects (as indicated by longitudinal studies in CF (9)). Regardless, these findings suggest that factors
307 intrinsic to airways predominate in establishing the early microbiota in these chronic respiratory infections.
308 Several observations support these conclusions. Culture-based studies of children with CF
309 demonstrated that radiographic lung disease, respiratory symptoms, and inflammation all precede the detection
310 of “standard CF pathogens” (14, 15, 31, 32). All of the shared core microbes identified in the current study, many
311 of which would traditionally be considered “oral flora”, were also commonly identified in oropharyngeal swabs
312 from infants with CF by pyrosequencing (33). The identification of such “oral flora” has been associated with both
313 inflammation (32) and structural lung disease (15) in children with CF. Similarly, piglets with engineered CFTR
314 mutations (34) concurrently develop diverse airway microbiota and lung disease similar to those observed in
315 children with CF. The cultured microbiota of CF pigs were similar in constituency to those of wild-type pigs, but
316 more persistent and higher in abundance. Therefore, the requirement for traditional pathogens for CF lung
317 disease or symptom development is questionable.
318 Both cross-sectional and longitudinal studies of CF respiratory samples using molecular methods have
319 demonstrated correlations between decreasing microbiota diversity, age, and disease progression (3, 8, 9, 30).
320 Interestingly, some organisms traditionally associated with chronic airways disease were classified as core
321 species in some disease cohorts but not others in our study (for example, *P. aeruginosa, S. aureus*, and several
322 species each of *Veillonella* and *Prevotella* were core members in both the CF and BE metacommunities, but
323 satellite for PBB, Table S1). Collectively, these results could indicate that particular microbes are associated with
324 and/or accelerate specific lung diseases. Alternatively, changes over time in airway microbiota, including the
325 gradual decrease in community diversity observed in CF and in COPD (28), and the gradual identification of *P.*

326 *aeruginosa* in people with CF and non-CF bronchiectasis (23), could all result from intensified antibiotic
327 treatment and/or underlying disease, but this concept would not require the microbes that eventually dominate to
328 be particularly pathogenic. Similarly, changes in the respiratory microbiota could result from evolution (therapy-
329 driven or otherwise) in another reservoir that seeds the airways, such as the oropharynx (as suggested by
330 abundance in our study of microbes traditionally associated with this space (35)) or the GI tract (which a recent
331 longitudinal study of infants with CF suggested could be a source for respiratory microbiota (33)).
332 Together with the above earlier studies, our results support a unified model of early airway infections, in
333 which a defect in the clearance of otherwise normal airway microbiota (for example, due to early airway injury,
334 altered mucus, defective ciliary motility, mechanical obstruction, or immunodeficiency) contributes to progressive
335 inflammation and airway damage. In this model, an extension of the "vicious cycle hypothesis" (13), either
336 medical treatment or the physical effects of underlying disease (such as diseased tissue or accumulated
337 secretions) could lead to progressive change in the microbiota. These microbiological changes (including
338 eventual selection for *P. aeruginosa* in CF or BE) could in turn accelerate lung disease, or they could be
339 consequences rather than causes of this progression. Future studies, either interventional or longitudinal, may
340 be able to refine this model.
341 Antibiotics play important roles in the management of all three of these symptomatic chronic infections.
342 For example, inhaled and oral antibiotics are commonly used in CF and BE both for maintenance (11, 36) and
343 exacerbation treatments (11, 13). For PBB, a prolonged course of antibiotics usually leads to resolution of
344 symptoms (18). These observations underscore the importance of bacteria in the pathogenesis of each disease;
345 the current findings provide a rationale for reexamining and comparing the culture-based antibiotic approaches
346 to each of these airway infections, particularly during early stages.
347 Our study was limited by the small sizes of each subject group. The groups were not matched for age,
348 treatment, disease severity, or demographics, limiting our ability to determine relationships between microbiota
349 and these factors. However, the similarity in microbiota among demographically and clinically divergent pediatric
350 groups could be interpreted to further support the concept of a common early pediatric respiratory microbiota;
351 additional study with demographically well-matched cohorts could strengthen this argument. We did not
352 distinguish between viable and nonviable cells in the current analysis due to the small volumes of many of the
353 samples analyzed, and thus the detection of a microbe cannot definitively be said to indicate its persistence
354 (although even transient or nonviable bacterial cells could potentially contribute to airway inflammation).

Similarly, it could be proposed that the microbial diversity found in our study and others is due to contamination by upper airway bacteria during sample collection. However, the significant similarity in microbiota among sputum and BAL samples from the children with BE in our study, as observed in studies of children with CF (5, 22), as well as the high abundance of microbes in our samples, argue against this issue significantly altering our findings.

In conclusion, children with early chronic lung infections due to three distinct diseases shared strikingly similar core airway microbiota that included many bacteria not generally identified by clinical laboratory culture. These similarities were independent of underlying diagnosis or disease severity, geography, antibiotic and steroid use, and age. These results indicate that different chronic airway infections begin similarly, remaining similar even after symptoms have begun, but diverging by adulthood. Our findings provide a baseline to compare with in future, longitudinal microbiota studies of diverse chronic airway infections, which would provide vital insight into the microbial determinants of chronic lung disease progression and could lead to improved treatments.

Acknowledgements

We thank the participating patients and their families, B. van Yserloo and W. Osborne for technical assistance, and R. Gibson and H. Smith-Vaughan for helpful discussions. This study was supported by grants from the Cystic Fibrosis Foundation (HOFFMA07P0), the NIH (K02HL105543) and the American Thoracic Society (CF-07-003) to LRH, UK Natural Environment Research Council (NE/H019456/1) to CJvdG. The studies in Brisbane were supported by Australia's National Health and Medical Research Council grants (1042601, 1040830, 1019834 and 1034703).

Data deposition: The sequence data reported in this paper have been submitted to the NCBI Short Read Archive database (Bioproject accession number: PRJNA200702).

References

1. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 2011;184:957–963.
2. Han MK, Huang YJ, Lipuma JJ, Boushey HA, Boucher RC, Cookson WO, Curtis JL, Erb-Downward J, Lynch SV, Sethi S, Toews GB, Young VB, Wolfgang MC, Huffnagle GB, Martinez FJ. Significance of the microbiome in obstructive lung disease. *Thorax* 2012;67:456–463.
3. Blainey PC, Milla CE, Cornfield DN, Quake SR. Quantitative analysis of the human airway microbial ecology reveals a pervasive signature for cystic fibrosis. *Sci Transl Med* 2012;4:153ra130.
4. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM, Lung HIV Microbiome Project. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *Am J Respir Crit Care Med* 2013;187:1067–1075.
5. Rogers GB, Carroll MP, Serisier DJ, Hockey PM, Jones G, Bruce KD. characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16s ribosomal DNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* 2004;42:5176–5183.
6. Cox MJ, Allgaier M, Taylor B, Baek MS, Huang YJ, Daly RA, Karaoz U, Andersen GL, Brown R, Fujimura KE, Wu B, Tran D, Koff J, Kleinhenz ME, Nielson D, Brodie EL, Lynch SV. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. *PLoS ONE* 2010;5:e11044.
7. Van der Gast CJ, Walker AW, Stressmann FA, Rogers GB, Scott P, Daniels TW, Carroll MP, Parkhill J, Bruce KD. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J* 2011;5:780–791.
8. Stressmann FA, Rogers GB, van der Gast CJ, Marsh P, Vermeer LS, Carroll MP, Hoffman L, Daniels TWV, Patel N, Forbes B, Bruce KD. Long-term cultivation-independent microbial diversity analysis demonstrates that bacterial communities infecting the adult cystic fibrosis lung show stability and resilience. *Thorax* 2012;doi:10.1136/thoraxjnl-2011-200932.
9. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, Cavalcoli JD, VanDevanter DR, Murray S, Li JZ, Young VB, LiPuma JJ. Decade-long bacterial community dynamics in cystic fibrosis airways. *Proc Natl Acad Sci USA* 2012;109:5809–5814.

10. Chang AB, Redding GJ, Everard ML. Chronic wet cough: Protracted bronchitis, chronic suppurative lung disease and bronchiectasis. *Pediatr Pulmonol* 2008;43:519–531.
11. Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am J Respir Crit Care Med* 2003;168:918–951.
12. Redding GJ. Bronchiectasis in children. *Pediatr Clin North Am* 2009;56:157–171, xi.
13. Chalmers JD, Smith MP, McHugh BJ, Doherty C, Govan JR, Hill AT. Short- and long-term antibiotic treatment reduces airway and systemic inflammation in non-cystic fibrosis bronchiectasis. *Am J Respir Crit Care Med* 2012;186:657–665.
14. Sagel SD, Gibson RL, Emerson J, McNamara S, Burns JL, Wagener JS, Ramsey BW. Impact of *Pseudomonas* and *Staphylococcus* infection on inflammation and clinical status in young children with cystic fibrosis. *J Pediatr* 2009;154:183–188.
15. Mott LS, Park J, Murray CP, Gangell CL, de Klerk NH, Robinson PJ, Robertson CF, Ranganathan SC, Sly PD, Stick SM. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 2012;67:509–516.
16. Rogers GB, Cuthbertson L, Hoffman LR, Wing PAC, Pope C, Hooftman DAP, Lilley AK, Oliver A, Carroll MP, Bruce KD, van der Gast CJ. Reducing bias in bacterial community analysis of lower respiratory infections. *ISME J* 2013;7:697–706.
17. Rogers GB, van der Gast CJ, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, Serisier DJ. Clinical measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax* 2013;68:731–737.
18. Chang AB, Yerkovich ST, Gibson PG, Anderson-James S, Petsky HL, Carroll ML, Masters IB, Marchant JM, Wurzel D, Upham JW. Pulmonary innate immunity in children with protracted bacterial bronchitis. *J Pediatr* 2012;161:621–625.e1.
19. Kapur N, Grimwood K, Masters IB, Morris PS, Chang AB. Lower airway microbiology and cellularity in children with newly diagnosed non-CF bronchiectasis. *Pediatr Pulmonol* 2012;47:300–307.
20. Wolter DJ, Emerson JC, McNamara S, Buccat AM, Qin X, Cochrane E, Houston LS, Rogers GB, Marsh P, Prehar K, Pope CE, Blackledge M, Déziel E, Bruce KD, Ramsey BW, Gibson RL, Burns JL, Hoffman LR. *Staphylococcus aureus* small-colony variants are independently associated with worse lung disease in children with cystic fibrosis. *Clin Infect Dis* 2013;doi:10.1093/cid/cit270.

21. Dowd SE, Wolcott RD, Sun Y, McKeehan T, Smith E, Rhoads D. Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). *PLoS ONE* 2008;3:e3326.
22. Harris JK, De Groote MA, Sagel SD, Zemanick ET, Kapsner R, Penvari C, Kaess H, Deterding RR, Accurso FJ, Pace NR. Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic fibrosis. *Proc Natl Acad Sci USA* 2007;104:20529–20533.
23. Maughan H, Cunningham KS, Wang PW, Zhang Y, Cypel M, Chaparro C, Tullis DE, Waddell TK, Keshavjee S, Liu M, Guttman DS, Hwang DM. Pulmonary bacterial communities in surgically resected noncystic fibrosis bronchiectasis lungs are similar to those in cystic fibrosis. *Pulm Med* 2012;2012:746358.
24. Bittar F, Richet H, Dubus J-C, Reynaud-Gaubert M, Stremler N, Sarles J, Raoult D, Rolain J-M. Molecular detection of multiple emerging pathogens in sputa from cystic fibrosis patients. *PLoS ONE* 2008;3:e2908.
25. Duff RM, Simmonds NJ, Davies JC, Wilson R, Alton EW, Pantelidis P, Cox MJ, Cookson WOCM, Bilton D, Moffatt MF. A molecular comparison of microbial communities in bronchiectasis and cystic fibrosis. *Eur Respir J* 2013;41:991–993.
26. Tunney MM, Einarsson GG, Wei L, Drain M, Klem ER, Cardwell C, Ennis M, Boucher RC, Wolfgang MC, Elborn JS. The Lung Microbiota and Bacterial Abundance in Patients With Bronchiectasis When Clinically Stable and During Exacerbation. *Am J Respir Crit Care Med* 2013;doi:10.1164/rccm.201210-1937OC.
27. Huang YJ, Nelson CE, Brodie EL, Desantis TZ, Baek MS, Liu J, Woyke T, Allgaier M, Bristow J, Wiener-Kronish JP, Sutherland ER, King TS, Icitovic N, Martin RJ, Calhoun WJ, Castro M, Denlinger LC, Dimango E, Kraft M, Peters SP, Wasserman SI, Wechsler ME, Boushey HA, Lynch SV. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. *J Allergy Clin Immunol* 2011;127:372–381.e1–3.
28. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB. Analysis of the lung microbiome in the “healthy” smoker and in COPD. *PLoS ONE* 2011;6:e16384.
29. Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, Cooper J, Sin DD, Mohn WW, Hogg JC. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;185:1073–1080.

30. Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, Wolfgang MC. The adult cystic fibrosis airway microbiota is stable over time and infection type, and highly resilient to antibiotic treatment of exacerbations. *PLoS ONE* 2012;7:e45001.
31. Zemanick ET, Wagner BD, Sagel SD, Stevens MJ, Accurso FJ, Harris JK. Reliability of quantitative real-time PCR for bacterial detection in cystic fibrosis airway specimens. *PLoS ONE* 2010;5:e15101.
32. Gangell C, Gard S, Douglas T, Park J, de Klerk N, Keil T, Brennan S, Ranganathan S, Robins-Browne R, Sly PD. Inflammatory responses to individual microorganisms in the lungs of children with cystic fibrosis. *Clin Infect Dis* 2011;53:425–432.
33. Madan JC, Koestler DC, Stanton BA, Davidson L, Moulton LA, Housman ML, Moore JH, Guill MF, Morrison HG, Sogin ML, Hampton TH, Karagas MR, Palumbo PE, Foster JA, Hibberd PL, O'Toole GA. Serial analysis of the gut and respiratory microbiome in cystic fibrosis in infancy: interaction between intestinal and respiratory tracts and impact of nutritional exposures. *MBio* 2012;3:
34. Stoltz DA, Meyerholz DK, Pezzulo AA, Ramachandran S, Rogan MP, Davis GJ, Hanfland RA, Wohlford-Lenane C, Dohrn CL, Bartlett JA, Nelson GA 4th, Chang EH, Taft PJ, Ludwig PS, Estin M, Hornick EE, Launspach JL, Samuel M, Rokhlina T, Karp PH, Ostedgaard LS, Uc A, Starner TD, Horswill AR, Brogden KA, Prather RS, Richter SS, Shilyansky J, McCray PB Jr, *et al.* Cystic fibrosis pigs develop lung disease and exhibit defective bacterial eradication at birth. *Sci Transl Med* 2010;2:29ra31.
35. Cardenas PA, Cooper PJ, Cox MJ, Chico M, Arias C, Moffatt MF, Cookson WO. Upper airways microbiota in antibiotic-naïve wheezing and healthy infants from the tropics of rural Ecuador. *PLoS ONE* 2012;7:e46803.
36. Barker AF, Couch L, Fiel SB, Gotfried MH, Ilowite J, Meyer KC, O'Donnell A, Sahn SA, Smith LJ, Stewart JO, Abuan T, Tully H, Van Dalfsen J, Wells CD, Quan J. Tobramycin solution for inhalation reduces sputum *Pseudomonas aeruginosa* density in bronchiectasis. *Am J Respir Crit Care Med* 2000;162:481–485.

493
494 **Figure legends**

495 **Fig. 1.** Box plot comparisons of bacterial diversity between the BE, PBB, and CF cohorts. Given are three
496 measures of diversity: species richness (*S**), Simpson's complement index (1-*D*), and Shannon-Wiener index
497 (*H*'). The top and bottom boundaries of each box indicate the 75^{th} and 25^{th} quartile values, respectively, and lines
498 within each box represent the 50^{th} quartile (median) values. Ends of whiskers mark the lowest and highest
499 diversity values in each instance. One-way ANOVA summary statistics are given in Table S2. The only
500 significant difference found was for *H'* between the CF and PBB cohorts (P=0.027, Table S2).
501
502 **Fig. 2**. Dendrograms of bacterial community membership for all patients from the BE, PBB, and CF cohorts.
503 Black, grey, and white shaded boxes are given to show positions of BE, PBB, and CF samples, respectively.
504 Patient species profiles were compared using the Sørensen index of similarity and un-weighted pair-group
505 averages (UPGMA).
506
507 **Fig. 3.** Distribution and dispersal of bacterial species among BE, PBB, and CF pediatric cohorts. (A) The number
508 of samples for which each detected bacterial species (open circles) was observed, plotted against the
509 abundance ($\log_{10}$ scale) of that species among all samples within each cohort (BE, $r^2 = 0.64$, $F_{1,175} = 311.5$, $P <$
510 0.0001; PBB, $r^2 = 0.72$, $F_{1,128} = 321.1$, $P < 0.0001$; and CF, $r^2 = 0.75$, $F_{1,141} = 418.4$, $P < 0.0001$). (B) A dispersal
511 plot to identify which bacterial species are randomly distributed within each cohort, a measure used to assign
512 core versus satellite status. Index of dispersion was calculated as the ratio of variance to mean of abundance for
513 each species within each cohort and plotted for each sample. The line depicts the 2.5 % confidence limit for the
514 χ^2 distribution. Species that fall below this line are randomly distributed and were considered satellite species,
515 whereas those that are above the line are non-randomly distributed and were considered core species. The 97.5
516 % confidence limit was not plotted, as no species fell below that line.
517
518 **Fig. 4.** Dendrograms of community membership similarity between the pediatric BE, PBB, and CF bacterial
519 metacommunities and compared with adult CF and BE metacommunities. Given are whole, core and satellite
520 microbiota. Metacommunity profiles were compared using the Sørensen index of similarity and UPGMA.
521 Similarities between the microbiota from different cohorts are read as the location of the horizontal line

522 connecting those cohorts (the “node”) on the Y-axis; for example, the similarity between pediatric core microbiota
523 was ≥ 90%, while that between adult core microbiota was ≥ 40%.
524
525 **Fig. 5.** Dendrograms of Raup and Crick (S_{RC}) probability-based index of similarity between the pediatric BE,
526 PBB, and CF bacterial metacommunities and compared with adult CF and BE metacommunities. Given are the
527 whole, core and satellite microbiota. $S_{RC} < 0.95$ and $S_{RC} > 0.05$ denote similarity no greater than expected by
528 chance. $S_{RC} < 0.05$ denotes significant dissimilarity and $S_{RC} > 0.95$ significant similarity. The 0.05 and 0.95
529 thresholds are depicted with dashed lines in each instance. Dendrograms were constructed using UPGMA.
530
531

Tables:

Table 1. Demographic characteristics of the pediatric study groups.

Clinical Characteristic	BE[1]	PBB	CF
Number of subjects	19	12	25
Geographic origin	Australia (Brisbane)	Australia (Brisbane)	US (Seattle)
Sample type	9 BAL & 10 Sputum	All BAL	All Sputum
Age (y; Mean ± SD, range)	8.9± 4.7, 1.8-16.3[6]	2.3 ± 1.7, 0.6-5.5[4, 5]	12.5±3.5, 2.3-17.7
Gender (% female)	36%	25%	68%
FEV_1 % predicted (Mean ± SD, range, no. subjects with measurements)	85.7 ± 21%, 41-120%, 15 subjects	117 ± 1.4%, 116 and 118%, 2 subjects[4, 5]	77.9 ± 24%, 22.2-116.2%, 21 subjects
Number (%) of subjects on antibiotics at sampling[2]	8 of 19 (42%)	1 of 12 (8%)	7 of 25 (28%)
Number (%) of subjects on corticosteroids at sampling	6 of 19 (32%), one oral, 5 inhaled	2 of 12 (16%), all inhaled	11 of 25 (48%), all inhaled
Number (%) culture-positive for *P. aeruginosa*[3]	0	0	8 (34%)
Number (%) culture-positive for *S. aureus*[3]	2 (11%)	2 (16%)	22 (96%)
Number (%) culture-positive for *H. influenzae*[3]	4 (19%)	6 (50%)	7 (30%)

[1] Of the 10 children with BE with additional diagnoses available, 5/10 (50%) had idiopathic bronchiectasis, one had previously had an airway foreign body, one had primary ciliary dyskinesia, one had chronic aspiration due to a tracheoesophageal fistula, one had Mounier-Kuhn syndrome, and one had common variable immunodeficiency.

[2] Antibiotics for children with CF included tobramycin (inhaled), linezolid, levofloxacin, ciprofloxacin, trimethoprim-sulfamethoxazole, ticarcillin-clavulanate, and amoxicillin-clavulanate. Antibiotics for children with PBB included only erythromycin. Antibiotics for children with BE included erythromycin, azithromycin, roxithromycin, clarithromycin, amoxicillin-clavulanate, cefotaxime, ceftriaxone, and ticarcillin-clavulanate.

[3] Of 23 subjects with CF with culture data available; all subjects with PBB and BE had available culture data.

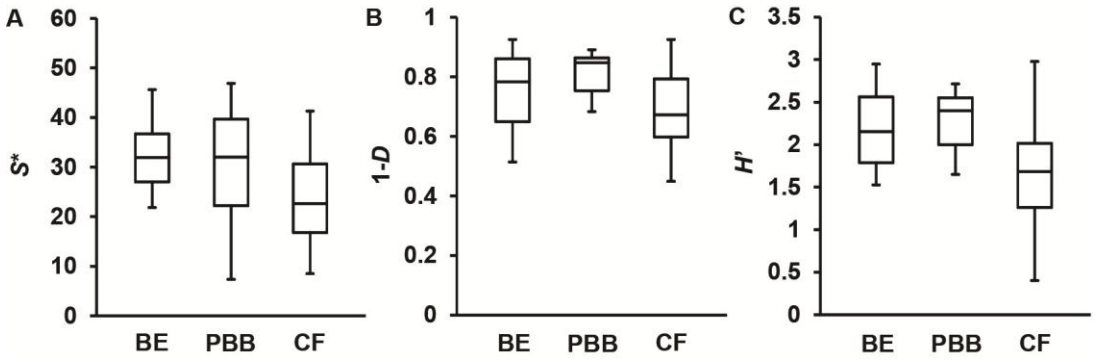
[4] $P < 0.001$ compared with CF; [5] $P < 0.001$ compared with BE; [6] $P < 0.01$ compared with CF.

Table 2. Similarity of Percentages (SIMPER) analysis of bacterial community similarity (Bray-Curtis) between pediatric metacommunities. Given is occupancy, or the number of patients for each disease group a given species was detected in. Next is mean % abundance of sequences for a species across the samples it was observed to occupy. Mean contribution represents the average contribution of a given species to the average similarity between samples (overall mean = 50.2%). Percentage contribution is the mean contribution divided by mean similarity across samples. The list of species is not exhaustive, so cumulative % value does not sum to 100%. Species level identities of detected taxa are reported here. However, given the length of the ribosomal sequences analysed, these identities should be considered putative.

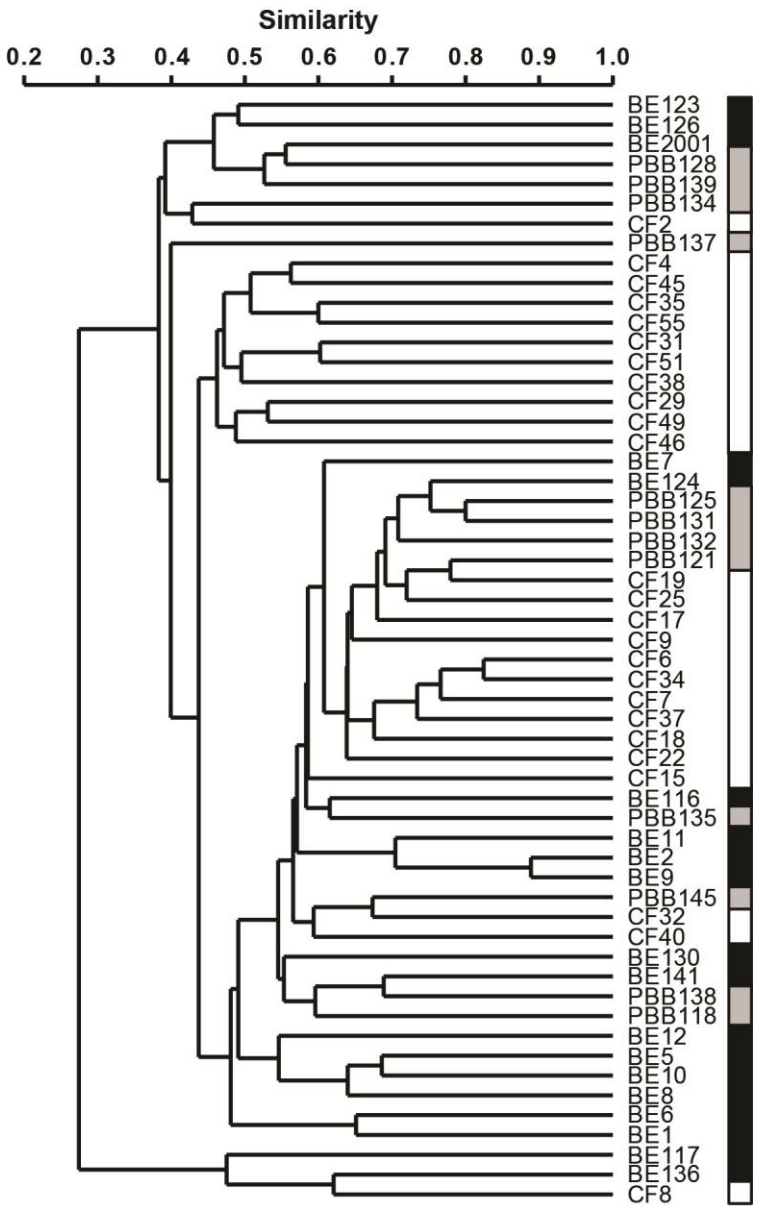
	Occupancy[1]						
Species	BE	PBB	CF	Mean Abund.	Mean Cont.	% Cont.	Cum. %
Haemophilus influenzae	12	11	19	23.22	18.79	37.44	37.44
Streptococcus mitis	18	12	25	14.59	8.45	16.84	54.28
Prevotella melaninogenica	16	11	24	9.16	4.59	9.15	63.43
Veillonella dispar	19	12	24	4.48	3.90	7.77	71.20
Fusobacterium nucleatum	17	11	21	2.63	1.99	3.97	75.17
Neisseria flavescens	12	11	19	5.04	1.73	3.44	78.61
Porphyromonas catoniae	17	11	18	1.66	1.05	2.09	80.70
Haemophilus parainfluenzae	15	10	16	1.30	1.04	2.06	82.76
Porphyromonas gingivalis	17	9	17	1.60	0.97	1.92	84.69
Prevotella nanceiensis	13	11	15	0.85	0.54	1.08	85.77
Prevotella histicola	12	7	19	0.81	0.46	0.92	86.69
Granulicatella adiacens	15	10	19	0.53	0.44	0.87	87.57
Prevotella oris	10	4	13	1.07	0.37	0.73	88.30
Sphingomonas echinoides	9	7	3	0.73	0.36	0.71	89.02
Gemella morbillorum	13	11	21	0.86	0.35	0.70	89.71
Prevotella pallens	16	8	13	0.63	0.32	0.63	90.35

[1]Total N for each pediatric disease cohort: BE, 19; PBB, 12; CF, 25.

Figure 1


A
S*
60
50
40
30
20
10
0
BE
PBB
CF
B
1-D
1
0.8
0.6
0.4
0.2
0
BE
PBB
CF
C
H'
3.5
3
2.5
2
1.5
1
0.5
0
BE
PBB
CF

582 Figure 2



583

584

585

586

587

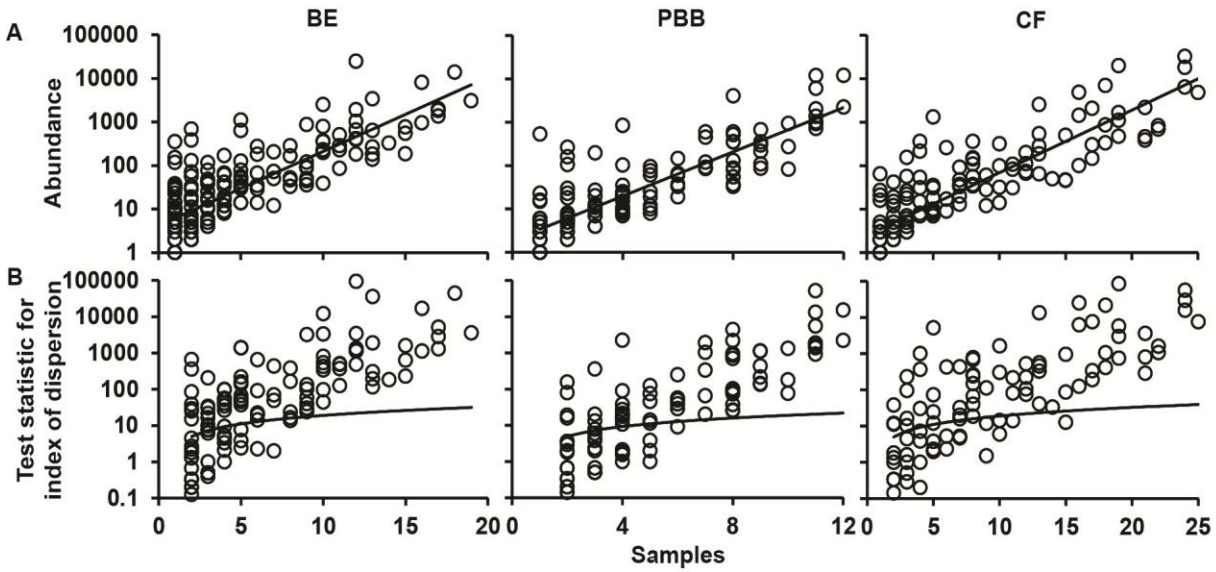
588

589

590

591

592 Figure 3



593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

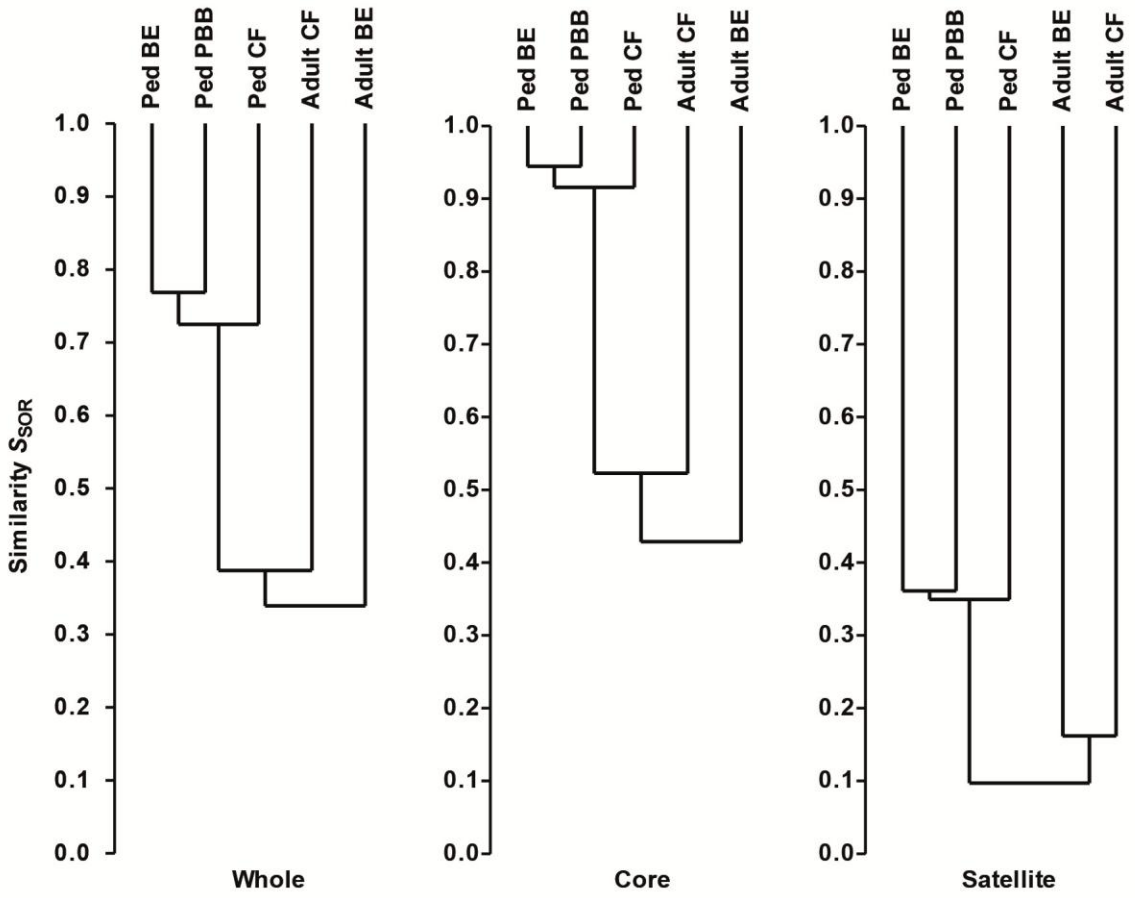
609

610

611

612

613 Figure 4



614

615

616

617

618

619

620

621

622

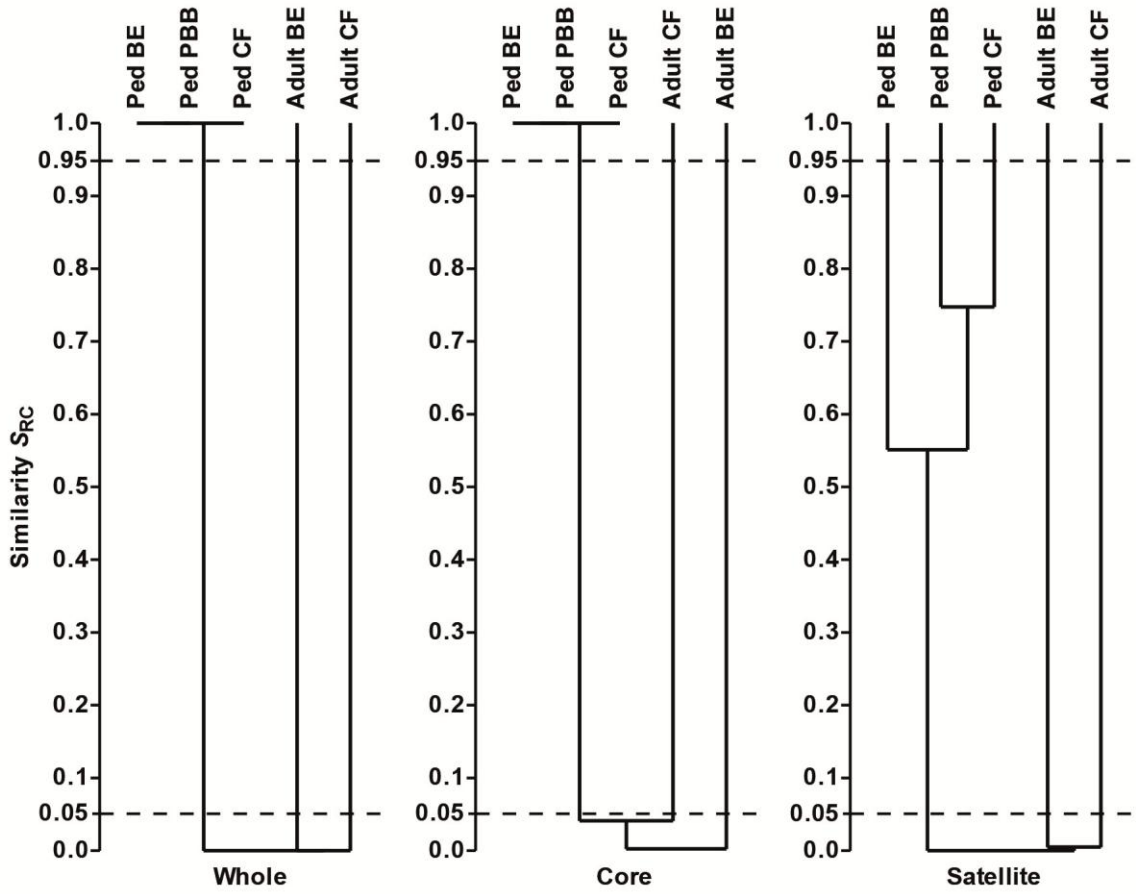
623

624

625

626

627 Figure 5



628

629